

## THE RIGHT NOT TO BE A FALSE POSITIVE: PROBLEMS IN THE APPLICATION OF THE DANGEROUSNESS STANDARD

Henry J. Steadman, Ph.D.

Among the rights of mental patients that have been affirmed in the past decade of mental health law activism, the right not to be a false positive is missing. This potential right has been overlooked, despite its close association with one of the major issues of this era—the definition and application of the dangerousness standard for involuntary commitment. That a right not to be a false positive was reasonable corollary to other protections surrounding the dangerousness standard was raised by Wilkins<sup>1</sup> in a rarely cited, but seminal article. Wilkins analyzed the moral trade-offs that society might be willing to make in its decisions for involuntary commitment or differential treatment because of assessments of a person's violence potential. Just how significant Wilkins' observation is becomes evident when the uses of the dangerousness standard for involuntary treatment and differential treatment are reviewed.

Shah<sup>2</sup> has cited at least 15 different points in the criminal justice and mental health systems where questions of an individual's dangerousness may be addressed. These range from decisions of emergency civil commitment in mental hospitals through release decisions from such facilities to the use of the death sentence in capital crime, as is the case currently in Texas. Fagin<sup>3</sup> noted that 38 of the 45 U.S. jurisdictions with emergency commitment statutes limit commitment to individuals who appear dangerous to themselves or others. The number of individuals retained under the dangerousness standard in state mental hospitals has been estimated by Scheidemandel and Kanno<sup>4</sup> to be about 50,000 people per year. An additional 30,000 persons per year are evaluated within the criminal justice system for dangerousness to decide either to which type of secure facility

---

Henry J. Steadman is Director of the Bureau of Special Projects Research New York State Department of Mental Hygiene Albany. Reprint requests should be addressed to Dr. Steadman at the New York State Department of Mental Hygiene, Albany, NY 12229. This paper will appear in *Psychiatric Patient Rights and Patient Advocacy: Issues and Evidence*, edited by Bernard C. Bloom and Shirley J. Asher to be published by Human Sciences Press. Permission to publish this paper in advance of its appearance in the book is gratefully acknowledged.

The research reported here was partially supported by PHS Grant MH28850 from the NIMH Center for Studies in Crime and Delinquency.

they will be sent or whether or not they will be released. Thus, estimations of a person's dangerousness are widespread in the United States in terms of both the numbers of persons and the range of situations to which they apply from child custody cases to release of defendants who have been acquitted by reason of insanity.

One of the basic difficulties in addressing the empirical evidence on predictions of dangerousness is imprecision as to what the concept of danger to self or others means conceptually and operationally. Megargee<sup>5</sup> has pointed out that "'dangerousness' is an unfortunate term, for it implies there is a trait of 'dangerousness' which, like intelligence, is a relatively constant characteristic of the person being assessed. . . . It is better to eschew the term 'dangerousness' in favor of discussing the problems involved in 'predicting dangerous behavior'" (p. 5). This distinction highlights two core characteristics of the concept of dangerousness. The first is that as a prediction, dangerousness is an estimation of the potential that a person will do something that is defined as dangerous. As such, dangerousness is a perception of the evaluator and not a characteristic, constant or otherwise, of the evaluatee. Second, dangerousness is by its nature a prediction. It means that because of certain characteristics or behaviors, a person is seen to have a high probability of performing certain acts in the future. Thus, the essence of dangerousness is that it is a perception and a prediction.

Because of its definition, one of the key questions in assessing the empirical data on predictions of dangerous behavior is the definition of dangerous behavior. Clearly dangerous behavior includes murder and assault and probably rape. However, the courts have included other behaviors such as writing "rubber" checks (*United States v. Charnizon* 232 A. 2d 586). Monahan's review<sup>6</sup> of the relevant research notes that the behaviors most often used to test predictive accuracy are arrest for violent crimes or physical assaults. In reviewing the empirical literature here both definitions will be employed, although it is more useful in mental health settings to limit dangerous behavior to any assaultive behavior.<sup>7</sup>

Before proceeding to review the empirical evidence on the accuracy of predicting dangerous behavior, it remains to clarify why Wilkins' discussion of the right not to be a false positive is so important. The dangerousness standard requires predictions by professionals within both the mental health and criminal justice systems, with psychiatrists the most frequent predictors. As Scheff<sup>8</sup> has observed, for all types of medical interventions, medical ideology favors consistently erring toward overtreatment, i.e., treating a healthy person rather than not treating the sick person. When this ideology is combined with infrequent behaviors, such as those included under the rubric of dangerous behavior, consistent overidentification of people as dangerous occurs. That is, in a group of 100 persons about, say, five might be expected to engage in an assaultive act in the next 12 months. In order to pick even three or four of these five who will be assaultive, with current levels of technology, probably 25 or 30 will be incorrectly identified. Public pressure and medical ideology both encourage the 25 or 30 to be incorrectly included in the assaultive group to control the three or four

who will be. Thus a false positive rate of eight or ten to one occurs. The alternative to reduce this false positive rate is to not identify any or at best a few of those who will actually be assaultive. Thus, what rights a mental patient or offender may have not to be so identified are important policy questions and are a productive framework in which to examine the empirical evidence on the accuracy of predictions of future dangerous behavior.

### EVIDENCE OF THE ACCURACY OF PREDICTIONS OF DANGEROUS BEHAVIOR

The reports dealing with the prediction of dangerous behavior fall into three general categories: (1) essays without data or with irrelevant data; (2) reviews of the primary research, occasionally with some secondary analysis; and (3) a few research studies with relevant primary data. The first category of articles predominate in law reviews and psychiatric journals. Most articles cited in the law journals and legal briefs have marginal data bases to them<sup>9,10</sup> or offer little more than anecdotes.<sup>11,12</sup> Many of these articles are seen as weighty contributions to the field and their legal views are no doubt important. If carefully examined, however, there is often more ideology than empiricism at the core of their arguments. The other type of article in this first category contains those that relate to the arrest rate of ex-mental patients. The inappropriate entry of these studies into assessments of predictive accuracy we have discussed elsewhere.<sup>13</sup> Since the majority of persons released from state mental hospitals are voluntary patients whose release in no way relates to psychiatric or legal assessments of their dangerousness, these arrest rate studies are at best tangentially related to questions of the validity of estimations of future assaultive behavior.

The second category of published work on dangerousness includes many valuable pieces. Among these works are those by Shah,<sup>2,14</sup> Monahan,<sup>15</sup> Laves,<sup>16</sup> Mesnikoff and Lauterbach,<sup>17</sup> and Rubin.<sup>18</sup> In many ways each of these reviews covers the same literature up to their respective publication dates and then draws some empirically grounded policy implications. In many ways these surveys are so comprehensive as to raise the question as to why this article was written. The answer to this question is twofold. First, there have been some significant developments in the evidence in this field even in the last year or two that warrant inclusion. Second, and perhaps more important, none of the prior reviews groups the research studies in a manner that highlights their key trends.

Of particular interest in this chapter is the distinction between the clinical and statistical predictions of dangerous behavior and the evidence on the various groups of professionals who are legally empowered to make such assessments in civil and criminal courts. It will become quite clear that, although this research literature is unusually consistent in its findings on both clinical and statistical prediction, there is considerably less evidence than there is generally assumed to be.

It is the third category of studies mentioned above that will be the focus here, i.e., those with primary data. Since many of these studies have been included in the review articles cited above, the earlier research will be skimmed with a greater concentration on four studies of more recent vintage together with the research and policy implications that are suggested by all the existing evidence.

### CLINICAL PREDICTION OF DANGEROUS BEHAVIOR

Most of the data related to the accuracy of clinical predictions of dangerous behavior come from research that followed mental patients released, contrary to psychiatric advice, from maximum security facilities by court decisions. The first study was that of Kozol and colleagues<sup>19</sup> that reported a follow-up of offenders released from the Bridgewater State Hospital program for dangerous sex offenders. One group (N=49) was released by the committing courts against the evaluation team's advice and another group (N=82) was released with psychiatric approval after diagnosis and treatment. The criterion for failure was arrest for a violent crime during the follow-up period. Of those released against psychiatric advice 35% were arrested for violent crime compared to 6% of those released after treatment with the approval of the psychiatric staff. One of the very serious problems with Kozol's work is that the length of time at risk of rearrest was not controlled.<sup>20</sup> Because those fully treated remained in Bridgewater, the group released against psychiatric advice could have been at risk as much as four years longer. Therefore, there are serious questions as to whether the two groups may be directly compared. What is clear is that even among the high recidivating group the false positive rate was about 2:1, with 35% accurately identified by the psychiatrists, but with 65% of those so identified not violently recidivating.

A second study which is often included in assessments of clinical accuracy in predicting dangerous behavior (e.g. Monahan<sup>15</sup>) is the 1973 report from Patuxent Institute in Maryland.<sup>21</sup> This report should not be included because of the widely differing treatment that the "comparison groups" received. As we have noted elsewhere<sup>22</sup> and as will be discussed below in more detail, all of the reports on the clinical successes of the Patuxent program published prior to 1977 compared "fully treated" individuals who had been in a very effective supervised parole program for up to three years with those released against the clinicians' advice whose recidivism rates began being calculated at the moment they first returned to the community. This meant that one group's recidivism was assessed only after they had been in the community for three years and all failures during this time were not counted. In contrast, the comparison group's failures in the same time period were counted. Thus, the methodological unsoundness precludes the inclusion of these Patuxent studies in the literature relevant to the questions under review here.

The next study relevant to clinical estimations of future dangerous

behavior was our follow-up of the Baxstrom patients.<sup>7</sup> This study followed 967 patients who had been inmates in maximum security correctional mental hospitals prior to the 1966 *Baxstrom v. Herold* decision of the United States Supreme Court. Following this ruling that appropriate due process protections had not been extended in the retention of this group, all 967 inmates were transferred en masse during a four month period to regular security mental hospitals despite having been retained for an average of 14 years because of their mental illness and dangerousness. During the four years that they were followed through mental hospitals and the community, 20% were assaultive at some time. Thus, the false positive rate was approximately four to one. Our contention that these transfers provide documentation for psychiatric inability to predict dangerous behavior accurately has been argued by others.<sup>23</sup> Given the information contained in the court decision, however, and the legal statutes under which they were detained, it seems that, indeed, the Baxstrom patient transfers were a naturalistic study of clinical predictions of dangerous behavior. Further discussion of the research and its critics can be found in Steadman and Cocozza.<sup>24</sup>

In an amazingly similar study, Thornberry and Jacoby<sup>25</sup> arrived at similar conclusions in Pennsylvania. Their study was strikingly similar in that they followed a group of 596 patients who were mass transfers after a Pennsylvania decision (*Dixon v. Attorney General of the Commonwealth of Pennsylvania*, 323, F. Supp. 966 (1971)) ruling that proper review had not occurred in decision of their retention in a hospital for the criminally insane. It is also similar in its findings. They found 14% of the 438 subjects at risk displayed some type of assaultive behavior in the community during a four year follow-up period. Thus, the false positive rate for an older, long-term institutionalized groups was about six to one.

The next study in the chronology of research on clinical predictions of dangerous behavior did not involve judicial intervention. In this work<sup>13</sup> the legal action precipitating the research was a statutory revision requiring two psychiatrists to assess whether indicted felony defendants who were incompetent to stand trial were dangerous. This determination by the court resulted in placement in either a mental hygiene facility, if not dangerous, or correctional facility, if dangerous. During the first year of this statute there were 257 males for whom a determination was made by psychiatrists. In 154 of these cases the defendants were evaluated as dangerous and in 113 they were evaluated as not dangerous. As reported in detail elsewhere,<sup>26</sup> 51% of the dangerous and 39% of the not dangerous were assaultive while hospitalized. After release, 16% of those evaluated as dangerous were assaultive resulting in either rehospitalization or arrest and 23% of the not dangerous group were assaultive. Overall, there was no difference in the frequency of assaultive behavior between the two groups beyond that obtainable by chance. Furthermore, although the clinical false positive rate for in-hospital assaultiveness was only 1:1, for community assaultiveness it was 5.4:1 (81 classified as dangerous who were not assaultive to 15 so classified who were assaultive).

Adding to the research on judicial interventions into psychiatric and administrative practices that produced clinical follow-up opportunities was the recently completed work reassessing the efficacy of the Patuxent Institute for Defective Delinquents in Maryland.<sup>22</sup> In this work, the comparison groups were reconceptualized from the in-house reports that had previously been reported. In our work, five groups were designated. They reflected all possible pathways into and through Patuxent. Four groups are of particular interest: the three evaluated as dangerous by the staff which include (1) a "fully treated" group, (2) a partially treated group released by the courts against staff recommendation, (3) a group disapproved for admissions despite staff estimations of their dangerousness, and (4) the group evaluated by the staff as not dangerous. Of the fully treated group 31% were arrested for violent crimes. Of the partially treated group 33% were arrested and of the group not admitted at all despite the staff's evaluations as dangerous 41% were arrested. Thus, the mean percentage arrested for violent crimes among the three study groups clinically evaluated as dangerous was 33.8%. However, in the fourth study group, those evaluated as not dangerous, 33.3% were arrested for violent offenses. There was no indication from these data of any ability on the part of the staff to identify accurately those who would be dangerous.

A study currently in progress in Texas directly addresses clinical accuracy in predicting dangerous behavior and is closely linked to judicial action. This study<sup>27</sup> involves the sequelae of a class action suit, *Renolds v. Neil*, that required the review of 188 inmates at Rusk State Hospital for the Criminally Insane for possible placement in less restrictive alternatives. To perform these evaluations the staff developed an instrument that was used in the assessments of all 188 inmates. Based on this instrument and other clinical information, 34 were defined as very dangerous and retained at Rusk, 87 were defined as dangerous thus requiring civil commitment, and 67 were discharged outright. Currently all three groups are being followed to determine the utility of the assessment instrument. Sheldon states that "To date, there has been no report of any serious criminal offense by any of the patients who were discharged with or without follow-up care." However, since the data are so incomplete at this time, any conclusions are premature.

The final study of clinical predictions involves no judicial action at all. It is also the only study of clinical predictions reviewed here in which the evaluations were not performed by psychiatrists. This research by Levinson and Ramsey<sup>28</sup> assessed the accuracy of predictions of dangerous behavior by paraprofessionals called mental health associates (MHA). Studying clients of a county emergency mental health unit, the files were checked to locate the routine estimates of danger to self or others that had been made during the work-ups. It was felt that since the MHAs were not as bound to the hospital and psychiatric ideology by training and job definition as were psychiatrists and because their backgrounds were closer to those of the clients than is typically the case with psychiatrists, they would have certain advantages in making such estimates. The data did not support his

## PSYCHIATRIC QUARTERLY

hypothesis. Considering only violent behavior, the MHA were wrong in 71% of the cases in which they predicted the person to be dangerous. However, the researchers found that there were substantial differences in the accuracy of the predictions based on perceptions of the level of stress in the clients' living situations. Where the environments were perceived as low stress, predictions were wrong in only seven of 26 cases (27%). Where the environments were seen as high stress they were wrong in 15 of 23 cases (65%). Nevertheless, the false positive rates did not vary. Instead, the increase in accuracy grew from the successful identification of the not dangerous group in the low stress settings. This means that the MHAs in this limited sample did have a false positive rate of 2.4:1 overall, which was better than most other clinical studies.

These, then, as summarized in the upper portion of Table 1 are the studies of clinical predictions of dangerous behavior. They are most consistent in that even among what are generally considered extremely high risk groups, clinical estimations rarely exceeded that which was obtainable simply by chance. Phrased another way, the predictive accuracy rarely exceeded the base rate of the behaviors predicted, i.e., where 40% accuracy is

TABLE 1  
False Positive Rates in Clinical and Statistical Studies  
of the Prediction of Dangerous Behavior

Clinical Predictions, N				
Study	Predicted Dangerous	Not Assaultive	Assaultive	False Positive Rate
Kozel et al.	49	32	17	1.9-1
Steadman and Cocozza	199	164	35	4.7-1
Thornberry and Jacoby	438	377	61	6.2-1
Steadman	257	170	87	2.0-1
Sheldon	121	—	—	—
Levinson and Ramsey	17	12	5	2.4-1
Steadman and Cocozza	154	75	79	0.95-1
	96	81	15	5.4-1
Statistical Predictions				
Wenk	—	—	—	6.1-1
Wenk	1,400	—	—	—
Hedland	138	83	55	1.5-1
Steadman and Cocozza	36	25	11	2.3-1
Jacoby	173	133	40	3.3-1
Koppin	60	31	29	1.1-1
Steadman and Cocozza	80	37	43	0.86-1
	51	36	15	2.4-1
	48	33	15	2.2-1

attained, about 40% of the total group for whom predictions were made exhibited the criterion behaviors. Thus clinical prediction attaining 40% accuracy would have been obtainable strictly by chance.

Even in those studies where the false positive rate was low, little special clinical acumen was apparent. Rather, it simply was that the base rate of the behavior in both the dangerous and not dangerous groups was so high that regardless of which group any individuals were placed in clinically, there was, for example, a 2:1 or in some cases an even chance that he would exhibit assaultive behavior. Thus, a low false positive rate reflects high base rate behavior rather than accurate clinical discriminations. Greater accuracy in each case was obtainable by predicting that no one would be dangerous. All other types of predictions increased the error rate, usually by identifying many persons as dangerous who were not on the indicators used. An excellent example of this phenomenon is the work of Bloom, Lang, and Goldberg<sup>29</sup> looking at clinical predictions of rehospitalization for mental patients. Of the 563 staff judgments for 92 released patients, 60% were correct in predicting no rehospitalizations within one year. However, of the 92 patients, 58% (53) remained out of hospital. Thus the predictions, although 60% accurate, were obtainable strictly by chance given the frequency of the criterion event. Thus, as we have noted elsewhere<sup>13</sup> if any of the evidential standards employed in criminal courts were applied to clinical predictions of dangerous behavior, none would be met, even the weakest of "more likely than not"; which is a probability of .51. There is simply no empirical evidence that psychiatrists or any other clinicians can clinically identify who will be dangerous beyond the accuracy anyone could attain simply by the probabilities of chance.

## STATISTICAL PREDICTIONS OF DANGEROUS BEHAVIOR

Given the inabilities of clinical predictions to show the level of accuracy needed to justify their expanding uses, as summarized in the lower section of Table 1, there have been a number of statistical forays into such predictions. A number of these have been linked to the research on clinical accuracy. The first set of very important studies by Wenk and co-workers<sup>30,31</sup> have been comprehensively and concisely reviewed by Monahan<sup>15</sup> and do not require further iteration here. It is sufficient to note that using sophisticated multivariate statistical procedures, Wenk was able to reduce his false positive rate among slightly over 4,000 California Youth Authority Wards to no less than 8:1. While his ratio was better than any that he felt were obtainable from the informal criteria employed by parole boards, which generally depended simply on a history of prior violence, the high statistical false positive rate precluded any direct application.

A second statistical prediction study that also was not linked to specific clinical predictions focused on 2,762 mental patients in mental hospitals in Missouri.<sup>32</sup> Using a very wide variety of sociodemographic, mental status, and admissions information, stepwise discriminant function analysis was



used in an attempt to identify those patients who were assaultive while hospitalized. On three criteria of assaultiveness the statistical "hits" were 90%, 90%, and 94%. However, the authors concluded that "they [including the prospective user of the predictive information] are faced with the inevitable dilemma of being wrong more often than right when a false positive prediction is made, even if we use the best predictive information available" (p. 446). This results from the low base rates of the respective behaviors which were 8%, 4%, and 10%. Thus, predictive accuracy could better be improved by always making negative predictions.

A third piece of statistical prediction was one segment of the Baxstrom research discusses above.<sup>7</sup> After determining a 4:1 false positive rate for the clinical predictions, we attempted to determine to what extent statistical prediction could have reduced this error rate. Of the various sociodemographic and criminal history variables available, the two that together best discriminated between those patients with assaults and those without assaults were age and the legal dangerousness scale (LDS), a summary scale of prior criminal history. The high risk group among the Baxstrom patients included those under 50 years of age and with an LDS score of 5 or more. This group included 80% of those who displayed assaultiveness in the community. Nevertheless, the false positive ratio still was 2:1 with two persons incorrectly identified as dangerous for everyone accurately designated.

The LDS scale was further tested at Colorado State Hospital on a group of released criminally insane patients. Koppin<sup>33</sup> employed a wide variety of psychiatric and social history indicators in conjunction with the LDS and obtained some statistically significant differentiations of those subsequently arrested for violence and those not. However, she concluded, "almost all accuracy rates computed were remarkably close to the base rate of 30% dangerous disruption among the patients in the sample" (p. 19). Similarly, when the LDS was employed in our study of incompetent felony defendants it was not as powerful an identifier, together with age, as had been the case among the Baxstrom patients. Only 22% of the high-risk group subsequently displayed violence producing a false positive rate of 3.6:1.

An analogous scale was developed by Pruesse and Quinsey.<sup>34</sup> Examining a group of 206 patients from the maximum security mental hospital at Penetanguishene, Ontario, a 0-5 point scale was developed to include such variables as the presence or absence of a diagnosis of personality disorder and length of time spent in mental hospitals. When the association between the scale and readmission was tested, it was statistically significant. When violent behavior was used as the dependent variable, however, there was no relationship.

The two other works relevant to the questions of statistical prediction of dangerous behavior both employed the statistical analysis used by Hedlund and co-workers<sup>32</sup> i.e., stepwise discriminant function analysis. The first is the work on the *Dixon* case discussed above.<sup>25,35</sup> They found that this type of analysis substantially reduced the level of false positives evident

in the clinical predictions. Their discriminant analysis correctly categorized 279 of the 432 patients as either not dangerous or dangerous. Assuming that the reason for the detention of the *Dixon* patients that they were deemed by psychiatrists and administrators as too dangerous, then only 64 of the 432 patients (15%) were identified accurately by clinicians, i.e., only 64 were subsequently violent. Thus, the clinical false positive rate was 5.5:1, whereas the discriminant function analysis classified 173 as dangerous of whom 40 were for a false positive rate of 3.3:1.

Recently this same type of analysis was applied to a group of incompetent felony defendants.<sup>36</sup> As seen in Table 2, in this sample also there was improvement over the clinical predictions. Working with a set of 53 possible independent variables ranging from height and weight through a history of drug or alcohol abuse, the discriminant analysis with the optimum prediction power for in-hospital assaults used seven factors. Of the 257 defendants, 80 were predicted to be dangerous. Of these, 43 were assaultive producing a false positive rate of .86:1, which means that for every 10 persons correctly identified 8.6 were incorrectly predicted to be dangerous. Of the 177 predicted not to be assaultive, 58 were indeed assaultive while hospitalized. Thus, the overall accuracy rate of the statistical prediction was 63%. This compares with the overall accuracy percentage of 46% by the psychiatrists and their false positive rate of .95:1.

When the same 53 variables were used to discriminate on the assaultiveness in the community, the statistical predictions were more impressive than for hospital violence. As seen in Table 3, the increase in predictive accuracy between the statistical and clinical predictions was moderate. Whereas the clinical predictions were accurate 59% of the time, 13 correctly identified as dangerous and 77 identified as not dangerous, the

TABLE 2  
Accuracy of Clinical and Statistical Predictions  
of Assaultiveness While Hospitalized

Actual groups	Clinical Prediction			
	Not Dangerous		Dangerous	
	N	%	N	%
Not assaultive	63	61.2	75	48.7
Assaultive	40	38.8	79	51.3
<i>Total</i>	103	100.0	154	100.0
	Statistical Prediction			
	N	%	N	%
Not assaultive	119	67.2	37	46.2
Assaultive	58	32.8	43	53.8
<i>Total</i>	177	100.0	80	100.0

TABLE 3  
Accuracy of Clinical and Statistical Analysis  
Predictions of Assaultiveness in the Community

Actual groups	Clinical Prediction			
	Not Dangerous		Dangerous	
	N	%	N	%
Not assaultive	77	86.5	50	79.4
Assaultive	12	13.5	13	20.6
<i>Total</i>	89	100.0	63	100.0

  

Actual groups	Statistical Prediction			
	Not Dangerous		Dangerous	
	N	%	N	%
Not assaultive	91	90.0	36	70.6
Assaultive	10	10.0	15	29.4
<i>Total</i>	101	100.0	51	100.0

statistical predictions were accurate in 70% of the cases with 10 persons incorrectly identified as not dangerous and 36 inaccurately predicted to be assaultive. Thus, the statistical false positive rate is 2.4:1 compared to the 3.8:1 clinical rate.

Much greater improvement is evident between the level of accuracy of statistical predictions and clinical ones in subsequent arrest of violent crimes as is seen in Table 4. Although the psychiatrists incorrectly predicted 57% of the cases, (83 were diagnosed as dangerous who were not

TABLE 4  
Accuracy of Clinical and Statistical Predictions of Subsequent Arrest for Murder, Manslaughter, or Assault

Actual groups	Clinical Prediction			
	Not Dangerous		Dangerous	
	N	%	N	%
No subsequent violent arrest	59	84.3	83	86.5
Some subsequent violent arrest	11	15.7	13	13.5
<i>Total</i>	70	100.0	96	100.0

  

Actual groups	Statistical Prediction			
	Not Dangerous		Dangerous	
	N	%	N	%
No subsequent violent arrest	109	92.4	33	68.8
Some subsequent violent arrest	9	7.6	15	31.2
<i>Total</i>	118	100.0	48	100.0

subsequently arrested for a violent crime and 11 were evaluated as not dangerous who were), the statistical predictions were inaccurate in only 25% of the cases (33 predicted to be dangerous who were not and 9 predicted to be not dangerous who were subsequently arrested for a violent crime). Thus, overall the statistical analysis correctly identified 124 of the 166 defendants and displayed a false positive rate of 2.2:1 (33 predicted wrongly to be dangerous and 15 correctly predicted to be dangerous) as compared to the false positive rate of 6.4:1 (83 to 13) experienced by the psychiatrists in their predictions.

In sum, it is clear from the statistical prediction studies reviewed that (1) in every case where comparisons were made, statistical prediction was superior to clinical prediction; (2) in most cases statistical predictions offered somewhat more accuracy than simple probabilities based on the base rates of the dangerous behaviors in questions; (3) in all cases the statistical predictions reduce the false positive rate of clinical predictions; and (4) the most accurate predictions of dangerous behavior remain those that say no one will be dangerous. Related to this fourth point, it must be noted that all clinical predictions analyzed were actual decisions about groups of patients thought to be unusually dangerous. On the other hand, none of the statistical studies were actually used in detention decisions. Thus, the latter predictions were unimpeded by the ever present and strong political pressures to err in a conservative direction by overpredicting who will be dangerous. Nevertheless, given the clear superiority of statistical prediction and inaccuracy of clinical predictions, there are serious questions about any patient's right not to be a false positive.

## IMPLICATIONS

It should be clear by this time that the research evidence on the prediction of dangerous behavior is consistent, but sparse. There is simply not that much evidence. What there is tends to be almost exclusively one type, follow-ups of groups for whom clinical predictions were made and who then spent substantial time institutionalized before some type of judicial intervention occurred requiring less restrictive settings. Monahan<sup>37</sup> has raised some probing questions about the adequacy of the available evidence to reach the conclusions that are being accepted about psychiatric inabilities to predict dangerous behavior. He suggests that some of the most important issues about clinical predictions relating to emergency commitment have yet to be addressed. Most of the research evidence, he argues, is not definitive because of the large amount of time between when predictions are made and the validating behaviors occur. Furthermore, in most of the research, even where the amount of time between the prediction and the follow-up is shorter, such as in the studies by Coccozza and Steadman<sup>38</sup> and Levinson and Ramsey,<sup>28</sup> treatment occurs. Thus, many serious gaps exist in the research evidence on clinical predictions.

Likewise, there has not been very extensive work in the statistical pre-

diction of dangerous behavior. What was done has been as consistent in its findings of improved accuracy over clinical predictions as have the clinical studies been consistent in showing little predictive expertise by clinicians. In every instance, overall the statistical predictions have been more accurate than the clinical prediction, particularly in reducing the false positive rates. This reduction is important not only for the moral and ethical issues which Wilkins<sup>7</sup> addresses, but also in terms of program costs. In many instances, evaluations of dangerousness result in placement in higher security facilities, which typically cost more to construct, have higher staff-patient ratios, and, as in the case in New York, have higher paid ward staff than the regular security facilities. Thus identification procedures which constantly overpredict, not only have implications in terms of patients rights, but also in terms of public expenditures. Nevertheless, as the data reviewed above indicate, there are severe restrictions in the ready application of the statistical predictions methods. These range from inherent limitations in accuracy, through the complexity of the statistical applications themselves, to limited testing and the ethics of detaining any individual because of statistical probabilities for groups into which his characteristics place him or her.

In examining the issues of predicting dangerous behavior and the application of the dangerousness standard it may be productive to turn the usual questions around. That is, rather than asking what evidence is there that psychiatrists, or other clinicians, cannot accurately predict dangerous behavior, what evidence is there that they *can*? When the question is phrased in this manner, the answer is unequivocal. *There is none.* Nowhere in the research literature is there any documentation that clinicians can predict dangerous behavior beyond the level of chance. Although there continue to be assertions of the viability of clinical judgment<sup>39</sup> that assure the listeners of accurate predictions and efficacious treatment to deter violence, there exists no empirical documentation. Clearly, as Monahan has pointed out, the range of research needed to assess the full range of relevant situations adequately has barely been tapped. Nevertheless, there is not a single piece of empirical evidence that accurate predictions under any circumstances are made by clinicians. There may be many instances in which they are quite accurate, but they have yet to demonstrate empirically they can. At this time it would seem appropriate to switch the burden of proof, given the consistency of the limited research evidence to the contrary.

It would appear that the legal activism that is demanding more from the predictors of dangerous behavior in the way of further specification of the factors that lead to their predictions, the time limits of their predictions, the behaviors that are being predicted, and how these behaviors seen as dangerous are logically derived from the clinical evidence is most appropriate. Monahan<sup>15</sup> for example, has argued that if one takes the 48-hour period that in many jurisdictions is the limit of emergency commitment, the clinician may be quite competent to make accurate assessments. This is, of course, an open question, but one for which there are no data. However,

when the predictions are extended to 60 days, or to months or years, it becomes meaningless in most instances. Likewise, if a person is considered dangerous because he deals in hard drugs, as was the case in one of the incompetent felony defendants in our research, and is about to be committed to a maximum security facility where drugs are not expected to be available for sale, he is not likely to be dangerous in that particular setting. Thus, the development of clear definitions of the behaviors to be entered as evidence as well as the specific behavioral expectations and the time frames of these predictions are important for due process protections in both the mental health and criminal justice systems.

This issue of due process protections is the key to understanding the predictions of dangerous behavior and the right not to be a false positive. Kittrie<sup>40</sup> has noted that dangerousness is a key concept of therapeutic state in that it masks, as is common with many treatment modalities, the actual use of the police power of the state under a *parens patriae* rationale for state intervention into an individual's life. The real issue in the commitment of the person for dangerousness is the state's justifiable right to protect its members. However, it is usually done as though it were in the best interest of the person committed, which it may also be. Because of this confusion between the rationales for involuntary commitment and the use of dangerousness, Wilkins' conceptualization of the right not to be a false positive becomes extremely important.

Since the use of predictions of dangerousness are really products of the state's right to protect its citizens, the question arises as to how often the state can be justified in detaining persons as dangerous who would not actually display the predicted behavior. That is, what is an acceptable false positive rate? That is, of course, a social policy question that frequently parades as a medical question of clinical judgment. Wilkins' suggests that this question should not be posed unilaterally. In this way, the moral trade-offs of inappropriate detention versus perceived needs for protection might be differentially applied to persons based on their history of prior violent behavior. In addition to any penalties of detention or fines, a price of criminal conviction, or documented assaults that resulted in hospitalization, would be an increase in the level of error that was acceptable. For example, if there is no history of violence, the level of error that would be tolerated might be none, as Wilkins suggests, or 5 in a 100, or the like. With one prior incident, this acceptable level might be 15 errors in 100 predictions. With multiple priors it would increase to 25 to 100 and so forth. As U.S. laws now stand, it would appear that the basic assumption is that no errors are being made or at least there are very few. Although criminal evidential standards as converted into mathematical probabilities<sup>13</sup> do allow some errors, these clearly depend on the seriousness of the resulting penalty. Thus the "more like than not" standard, or 51% level of certainty, is not an acceptable standard in a capital punishment case. In such instances the level of certainty must be "beyond a reasonable doubt" at about a 95 or 99% certainty. As yet the application of such varying evidentiary standards have not been discussed, let alone implemented in the area

of predicting dangerous behavior. Given the evidence presented here it is clear that such discussions are core clarifications needed in an area of muddled clinical and social policy debate. The evidence is limited, but consistent. False positive rates are high, greatly exceeding any accepted criminal law evidential standards. Whether or to what extent a person may have a right not to be a false positive is a question that clearly emerges from the data. Not only must a wider range of research be designed to address the scope of circumstances in which predictions of dangerous behavior are relevant, but also policy analyses must begin to demarcate the scope of patients' right not to be a false positive in the application of the dangerousness standard.

## REFERENCES

1. Wilkins LT: Current aspects of penology: Directions for corrections. *Proc Am Philosoph Soc* 118(3):235-247, 1974.
2. Shah SA: Dangerousness: A paradigm for exploring some issues in law and psychology. *Am Psychologist* March: 224-238, 1978.
3. Fagin A: The policy implications of predictive decision-making: "Likelihood" and "dangerousness" in civil commitment proceedings. *Public Policy* 24(4):491-528, 1976.
4. Scheidemandel PL, Kanno CK: The mentally ill offender: A survey of treatment programs. Washington, D.C., The Joint Information Service of the American Psychiatric Association and the National Association for Mental Health, 1969.
5. Megargee EI: The prediction of dangerous behavior. *Criminal Justice Behav* 3(1):3-22, 1976.
6. Monahan J: Social policy implications of the inability to predict violence. *J Soc Iss* 31(2):153-164, 1975.
7. Steadman HJ, Cocozza JJ: *Careers of the Criminally Insane*. Lexington, Mass., Lexington Books, 1974.
8. Scheff T: *Being Mentally Ill*. New York, Aldine, 1966.
9. Ennis BJ, Litwack TR: Psychiatry and the presumption of expertise: Flipping coins in the courtroom. *Calif Law Rev* 62:693-752, 1976.
10. Stone AA: *Mental Health and Law: A System in Transition*. Rockville, Md, National Institute of Mental Health, 1975.
11. Treffert D: Dying with your rights on. Paper presented at the Annual Meeting of the American Psychiatric Association, Detroit, May, 1974.
12. Peele R, Chadoff P, Taub N: Involuntary hospitalizations and treatability: Observations from the District of Columbia experience. *Cathol Univ Law Rev* 23:744-753, 1974.
13. Cocozza JJ, Steadman HJ: The failure of psychiatric predictions of dangerousness: Clear and convincing evidence. *Rutgers Law Rev* 29(5):1084-1101, 1976.
14. Shah SA: Dangerousness: Some definitional, conceptual, and public policy issues. In B Sales (ed.): *Perspectives in Law and Psychology*. New York, Pergamon, 1977.
15. Monahan J: The prediction of violent criminal behavior: A methodological critique and prospectus. In *Deterrence and Incapacities: Estimating the Effects of Criminal Sanctions on Crime Rates*. Washington, D.C., National Academy of Science, 1978.
16. Laves RG: The prediction of "dangerousness" as a criterion for involuntary civil commitment: Constitutional considerations. *J Psychiatry Law* 3(3):292-326, 1975.
17. Mesnikoff AM, Lauterbach CG: The association of violent dangerous behavior with psychiatric disorders: A review of the research literature. *J Psychiatry Law* 3(4):415-445, 1975.
18. Rubin B: Prediction of dangerousness in mentally ill criminals. *Arch Gen Psychiatry* 27:397-407, 1972.

19. Kozol H, Boucher R, Garolfalo R: The diagnosis and treatment of dangerousness. *Crime Delinquency* 18:371-392, 1972.
20. Cocozza JJ: Dangerousness. *Psychiatr News* 15:2, 1973 (August).
21. *Maryland's Defective Delinquency Statute—A Progress Report*. Department of Public Safety and Correctional Services. Unpublished manuscript. Baltimore, State of Maryland, 1973.
22. Steadman HJ: A new look at recidivism among patuxent inmates. *Bull Am Acad Psychiatry Law* 5(2):200-209, 1977.
23. Halpern AL: Review of careers of the criminally insane. *Bull Am Acad Psychiatry Law* 4(2):187-191, 1975.
24. Steadman HJ, Cocozza JJ: The prediction of dangerousness—*Baxstrom*: A case study. In G. Cooke (ed): *Readings in Forensic Psychology*. Springfield, Ill., Charles C. Thomas, forthcoming.
25. Thornberry TP, Jacoby JE: The uses of discretion in a maximum security mental hospital: The Dixon case. Paper presented at the Annual Meeting of the American Society of Criminology, Chicago, 1974.
26. Steadman HJ, Cocozza JJ: Psychiatry, dangerousness and the repetitively violent offender. *J Crim Law Criminol* 69(2):226-231, 1978.
27. Sheldon RB: Assessing dangerousness in the criminally insane. Paper presented at the American Psychological Association Meeting, San Francisco, 1977.
28. Levinson RM, Ramsay G: Dangerousness, stress and mental health evaluations. *J Health Soc Behav* in press.
29. Bloom BL, Lang EW, Goldberg H: Factors associated with accuracy of prediction of posthospitalization adjustment. *J Abnorm Psychol* 76(2):243-249, 1970.
30. Wenk E, Emrich R: Assaultive youth: An exploratory study of the assaultive experience and assaultive potential of California Youth Authority wards. *J Res Crime Delinqu* 9:171-196, 1972.
31. Wenk E, Robinson JO, Smith GW: Can violence be predicted? *Crime Delinqu* 18:393-402, 1972.
32. Hedlund L, Sletten IW, Altman H, Evenson RC: Prediction of patients who are dangerous to others. *J Clin Psychol* 29(4):443-447, 1973.
33. Koppin, M. *Age, hospital stay and criminal history as predictors of post-release release danger*. Pueblo, Colo., Colorado State Hospital, 1977.
34. Pruesse M, Quinsey VL: The dangerousness of patients released from maximum security: A replication. *J Psychiatry Law* 5(2):293-299, 1977.
35. Jacoby JE: Prediction of dangerousness among mentally ill offenders. Paper presented at Annual Meeting of the American Society of Criminology, Toronto, 1975.
36. Steadman HJ, Cocozza JJ: The dangerousness standard and psychiatry: A cross national issue in the social control of the mentally ill. Paper presented at the 9th World Congress of Sociology, Uppsala, Sweden, 1978.
37. Monahan J: Prediction research and the emergency commitment of dangerous mentally ill persons: A reconsideration. *Am J Psychiatry* 135(2):198-201, 1978.
38. Cocozza JJ, Steadman HJ: Prediction in psychiatry: An example of misplaced confidence in experts. *Social Problems* 25(3):265-276, 1978.
39. Kinzel A: Confronting and identifying dangerousness. *Am J Psychiatry* 132(12):1331, 1975.
40. Kittrie N: *The Right to Be Different*. Baltimore, Penguin Books, 1971.